Enfoque programático para optimizar la configuración de VPN de acceso remoto mediante análisis de datos

Contenido

Introducción

Este documento describe cómo monitorear y optimizar la VPN de acceso remoto configurada a través de algunos de los módulos de programación y herramientas de código abierto disponibles hoy en día. Actualmente se generan muchos datos incluso en las redes más pequeñas que se pueden utilizar para obtener información útil. La aplicación de análisis de estos datos recopilados ayuda a tomar decisiones empresariales más rápidas y mejor fundamentadas, respaldadas por hechos.

Prerequisites

Requirements

Cisco recomienda que tenga conocimiento sobre estos temas:

- VPN de acceso remoto
- Conceptos básicos de programación de Python

Componentes Utilizados

Este documento no se limita a versiones específicas de hardware y software de Cisco ASA o FTD.

> **Nota:** Pandas, Streamlit, CSV y Matplotlib son algunas bibliotecas Python que se utilizan.

The information in this document was created from the devices in a specific lab environment. All of

the devices used in this document started with a cleared (default) configuration. Si su red está activa, asegúrese de comprender el impacto potencial de cualquier comando y scripts python.

Problema

Como muchas empresas adoptan el modelo de trabajo desde casa para la mayoría de sus empleados en todo el mundo, el número de usuarios que confían en VPN para llevar a cabo su trabajo ha aumentado considerablemente. Esto ha llevado a un aumento repentino y considerable de la carga en los concentradores VPN, lo que ha llevado a los administradores a replantearse y replanificar sus configuraciones de VPN. La toma de decisiones informadas para reducir la carga en los concentradores ASA requiere recopilar una amplia gama de información de los dispositivos durante un período de tiempo y evaluar esa información, que es una tarea compleja y requeriría un tiempo considerable si se hiciera manualmente.

Solución

Con varios módulos de Python y herramientas de código abierto disponibles actualmente para la programación de la red y el análisis de datos, la programación puede resultar muy útil en la recopilación y el análisis de datos, la planificación y la optimización de la configuración de VPN.

Análisis inicial basado en usuarios VPN y conexiones simultáneas

Para iniciar el análisis, obtenga el número de usuarios que se conectan, las conexiones simultáneas establecidas y su impacto en el ancho de banda. Los siguientes resultados del comando Cisco ASA proporcionan estos detalles:

- **show vpn-sessiondb anyconnect**
- **show conn**

Módulo Python **Netmiko** se puede utilizar para enviar al dispositivo, ejecutar los comandos y analizar los resultados.

```
cisco_asa_device = {

       "host": host,

       "username": username,

       "password": password,

       "secret": secret,

       "device_type": "cisco_asa",

   }

   net_conn = ConnectHandler(**cisco_asa_device)

   command = "show vpn-sessiondb anyconnect"

   command_output = net_conn.send_command(command)
```

Recopile el número de usuarios de VPN y el número de conexiones a intervalos regulares (cada 2 horas puede ser un buen comienzo) en una lista y obtenga el recuento diario máximo de un día.

```
#list1 is the list of user counts collected in a day
#list2 is the list of connection counts in a day
list1.sort()
max_vpn_user = list1[-1]

list2.sort()
max_conn = list2[-1]

df1.append([max_vpn_user,max_conn])
```

Pandas es una biblioteca de análisis y manipulación de datos eficiente y todos los datos analizados se pueden almacenar como una serie o trama de datos en pandas, lo que facilita las operaciones con los datos.

```
import pandas as pd

df = pd.DataFrame(df1, columns=['Max Daily VPN Users Count','Max Daily Concurrent
Connections'],index=<date range>)
```

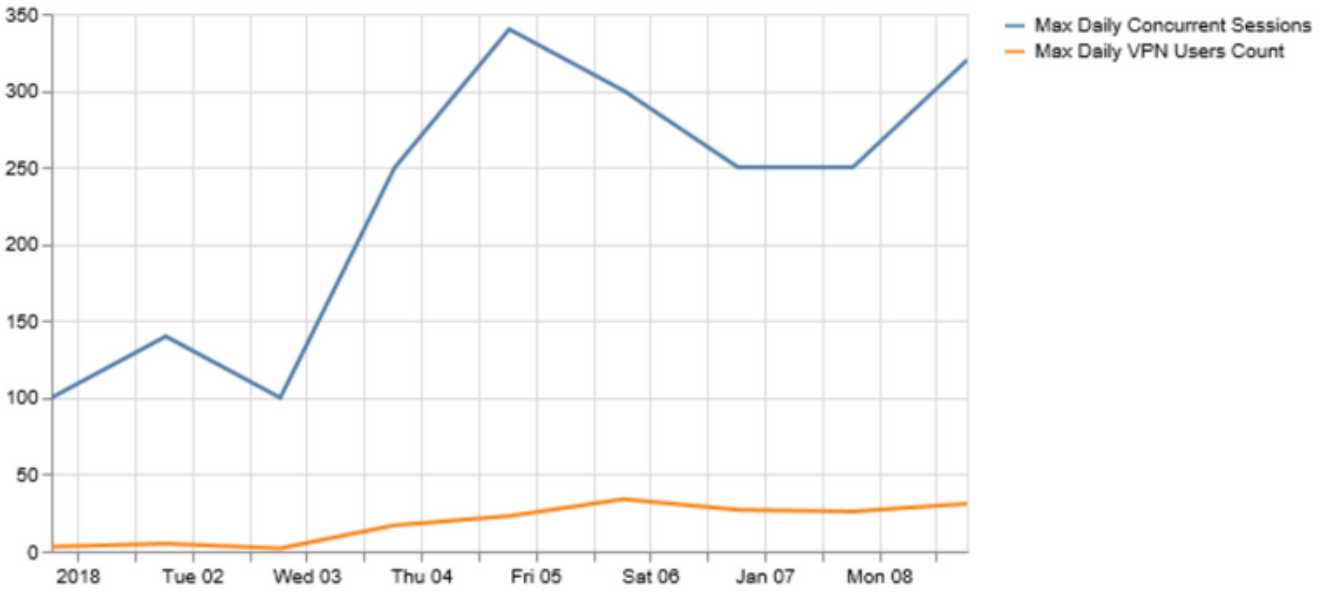
Daily Max VPN user Count - Max concurrent count

	Max Daily VPN Users Count	Max Daily Concurrent Sessions
Jan 1, 2018	3	100
Jan 2, 2018	5	140
Jan 3, 2018	2	100
Jan 4, 2018	17	250
Jan 5, 2018	23	340
Jan 6, 2018	34	300
Jan 7, 2018	27	250
Jan 8, 2018	26	250
Jan 9, 2018	31	320

Analice el **número máximo diario de usuarios de VPN** y **el número máximo de conexiones simultáneas** que pueden ayudar a determinar la necesidad de optimizar la configuración de VPN.

Utilice la función de trazado en pandas y biblioteca **matplotlib**, como se muestra en la imagen aquí.

```
df.plot()

matplotlib.pyplot.show()
```



Si el número de usuarios VPN o conexiones simultáneas se acerca a la capacidad de la cabecera VPN, puede causar estos problemas:

- Nuevos usuarios de VPN que se descartan.
- Se descartan nuevas conexiones de datos a través del ASA y los usuarios no pueden acceder a los recursos.
- Alta CPU y/o memoria.

La tendencia a lo largo de un período de tiempo puede ayudar a determinar si la caja está alcanzando su umbral.

Identificar la tendencia del tráfico hacia la red interna o hacia las redes externas

Show conn output en Cisco ASA puede proporcionar detalles adicionales tales como si el tráfico es hacia redes internas o externas y cuánta información en bytes por flujo pasa a través del firewall.

Soure IP	Destination IP	Service	Bytes
10.10.1.1	10.30.2.2	tcp/445	1234
10.10.1.2	40.5.2.3	tcp/443	2341
10.10.1.4	42.4.2.33	tcp/80	5432
10.10.2.3	52.3.2.34	tcp/443	1223
10.10.6.5	10.30.22.2	tcp/80	212
10.10.3.2	10.30.2.3	udp/389	1212
10.10.3.4	32.3.22.2	tcp/443	2123

El uso del módulo Netaddr python facilita la división de la tabla de conexión obtenida en flujos a redes externas y a redes internas.

```
for f in df['Responder IP']:
   private.append(IPAddress(f).is_private())

df['private'] = private

df_ext = df[df['private'] == False]

df_int = df[df[‘private’] == True]
```

Esta es la imagen del tráfico interno.

Soure IP	Destination	Service	Bytes
10.10.1.1	10.30.2.2	tcp/445	1234
10.10.6.5	10.30.22.2	tcp/80	212
10.10.3.2	10.30.2.3	udp/389	1212

Esta es la imagen del tráfico externo.

Soure IP	Destination	Service	Bytes
10.10.1.2	40.5.2.3	tcp/443	2341
10.10.1.4	42.4.2.33	tcp/80	5432
10.10.2.3	52.3.2.34	tcp/443	1223
10.10.3.4	32.3.22.2	tcp/443	2123

De este modo, proporciona una perspectiva del porcentaje de tráfico VPN destinado a las redes internas y de la cantidad de tráfico que sale a Internet. La recopilación de esta información durante un período de tiempo y el análisis de su tendencia pueden ayudar a determinar si el tráfico VPN es predominantemente externo o interno.

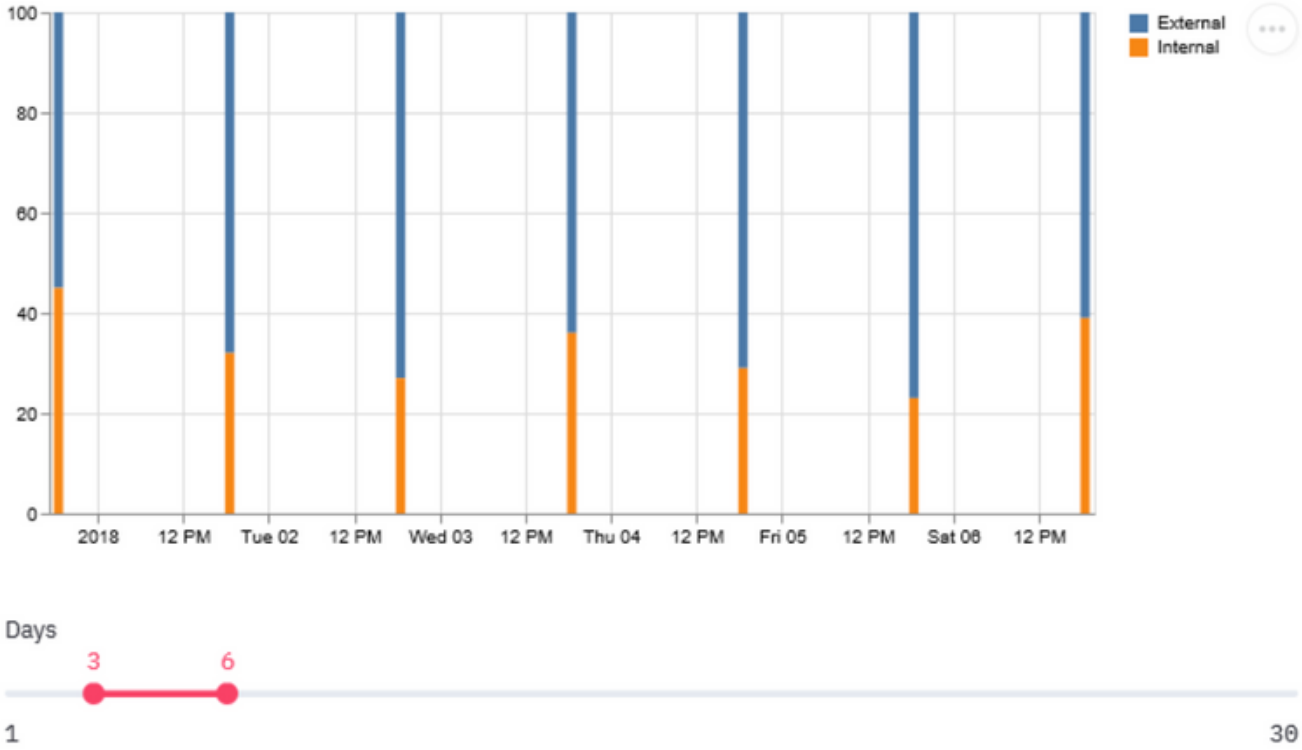
VPN Usage

Traffic Segregation - Internal and External

	External	Internal
Jan 1, 2018	55	45
Jan 2, 2018	68	32
Jan 3, 2018	73	27
Jan 4, 2018	64	36
Jan 5, 2018	71	29
Jan 6, 2018	77	23
Jan 7, 2018	61	39

Módulos como **Streamlit** permiten no sólo convertir los datos tabulares en una representación gráfica, sino también aplicar modificaciones en tiempo real para facilitar el análisis. Puede modificar la ventana de tiempo de los datos recopilados o agregar datos adicionales a los parámetros que se supervisan.

```
import streamlit

#traffic_ptg being a 2D array containing the data collected as in the table above

d = st.slider('Days',1,30,(1,7))

idx = pd.date_range('2018-01-01', periods=7, freq='D')

df = pd.DataFrame(d<subset of the list traffic_ptg based on slider
value>,columns=['External','Internal'],index=idx)

st.bar_chart(df)
```

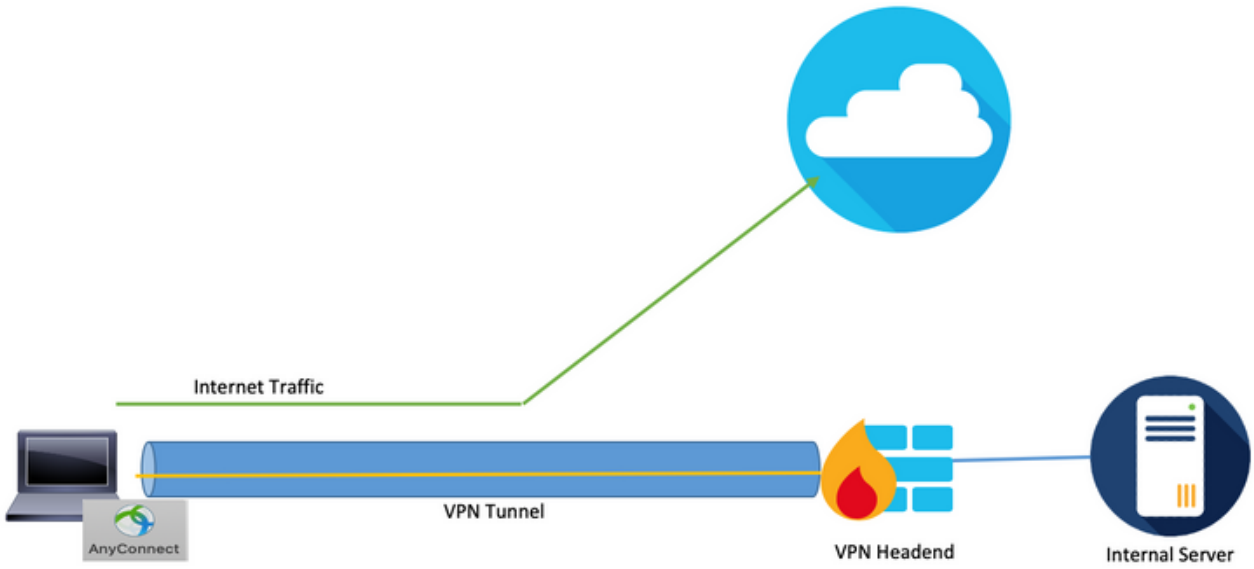


Una tendencia que se inclina hacia un mayor tráfico interno podría significar que la mayoría de los usuarios de VPN acceden a los recursos internos. Por lo tanto, para hacer frente a esto, aumentar la carga, es importante planificar actualizaciones en cajas más grandes o compartir la carga con conceptos como el balanceo de carga VPN.

En algunos casos, la capacidad de VPN puede estar todavía por debajo del umbral, pero un aumento en el número de usuarios de VPN puede agotar el conjunto de VPN configurado actualmente. En estos casos, aumente el VPN IP Pool.

Sin embargo, si la tendencia muestra que la mayoría del tráfico VPN es externo, puede utilizar la tunelización dividida.

Utilización de la función de tunelización dividida



Se trata de una función que envía sólo un conjunto específico de tráfico a través del túnel desde el sistema de usuario y el resto del tráfico se reenvía al gateway predeterminado sin cifrado VPN. Por lo tanto, para reducir la carga en el concentrador VPN, sólo el tráfico destinado a la red interna se puede rutear a través del túnel y el tráfico de Internet se puede reenviar a través del ISP local del usuario. Se trata de un método eficaz y ampliamente adoptado, pero conlleva algunos riesgos.

Un empleado accede a algunos sitios de redes sociales a través de redes no protegidas para una rápida interrupción, lo que puede infectar su portátil con malware que se propaga por toda la empresa debido a la falta de los niveles de seguridad de defensa en profundidad que se configuran en el lugar de trabajo. Una vez infectado, el dispositivo comprometido podría convertirse en un punto de inflexión desde Internet hacia el segmento de confianza, al pasar por alto las defensas perimetrales.

Una manera de reducir el riesgo mientras se utiliza esta función sería utilizar la tunelización dividida sólo para los servicios en la nube que superen criterios de seguridad estrictos, incluida una buena higiene de los datos y compatibilidad con Duo Security. La adopción de esta opción ayudará a aquellos servicios en la nube seguros que destinen una buena parte del tráfico externo observado anteriormente. Esto plantea la necesidad de analizar las aplicaciones web a las que acceden los usuarios de VPN.

La mayoría de los firewalls de última generación, como Cisco Firepower Threat Defense (FTD), contienen información de la aplicación asociada al evento en los registros. El análisis y la limpieza de estos datos de registro con **las bibliotecas csv** de python y las funciones de manipulación de datos de pandas pueden proporcionar un conjunto de datos similar al anterior con una adición de las aplicaciones a las que se está accediendo mapeadas.

```
#connections.csv contains the connection events from ASA and events_with_app.csv contains
connection events with Application details fromFTD

df1 = pd.read_csv(‘connections.csv’) df2 = pd.read_csv(‘events_with_app.csv') df_merged =
pd.merge(df1,df2,on=['Source IP','Destination IP','Service'])
```

Soure IP	Destination IP	Service	Bytes	Application
10.10.1.1	10.30.2.2	tcp/445	1234	
10.10.1.2	40.5.2.3	tcp/443	2341	Microsoft
10.10.1.4	42.4.2.33	tcp/80	5432	Microsoft
10.10.2.3	52.3.2.34	tcp/443	1223	Office365
10.10.6.5	10.30.22.2	tcp/80	212	
10.10.3.2	10.30.2.3	udp/389	1212	
10.10.3.4	32.3.22.2	tcp/443	2123	Youtube

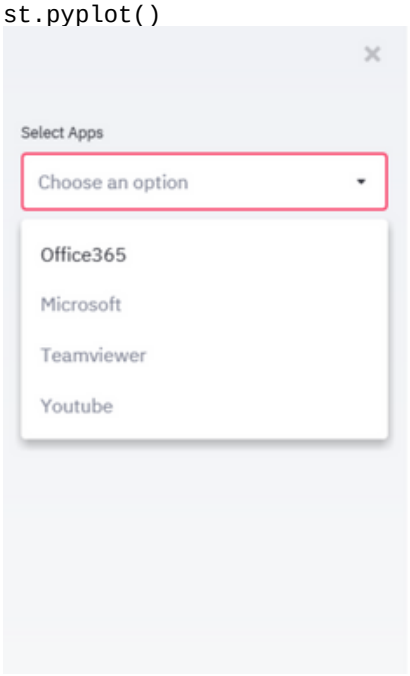
Una vez obtenida una trama de datos como la anterior, puede categorizar el tráfico externo total basado en la aplicación a través de pandas.

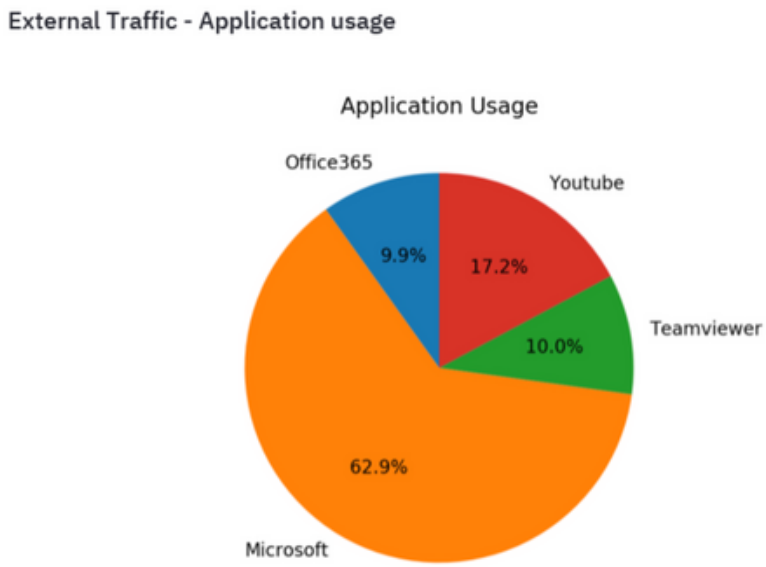
```
df2 = df.groupby('Application')

df3 = df2['Bytes'].sum()
```

```
Application
Microsoft     7773
Office365     1223
Teamviewer    1234
Youtube       2123
Name: Bytes, dtype: int64
```

El uso de Streamlit obtiene de nuevo una representación gráfica del porcentaje de cada aplicación en el tráfico total. Permite la flexibilidad de cambiar la ventana de tiempo para la inclusión de los datos, así como filtrar las aplicaciones en la propia interfaz de usuario sin necesidad de realizar cambios en el código, lo que hace que el análisis sea fácil y preciso.

```
import matplotlib.pyplot as plt

apps = ['Office365', 'Microsoft', 'Teamviewer', 'Youtube']
app_select = st.sidebar.multiselect('Select Apps',activities)

# app_bytes - list containing the applications and bytes

plt.pie(app_bytes, labels=apps)
plt.title('Application Usage')

st.pyplot()
```





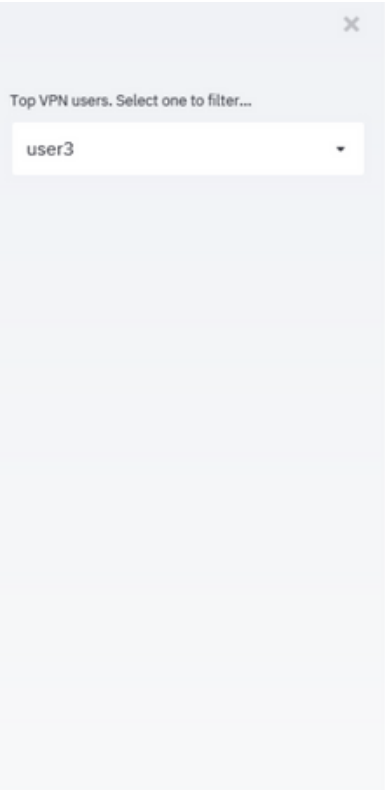
Esto puede simplificar el proceso de identificación de las principales aplicaciones web que utilizan los usuarios de VPN durante un período de tiempo y si estas aplicaciones sirven para proteger o no los servicios en la nube.

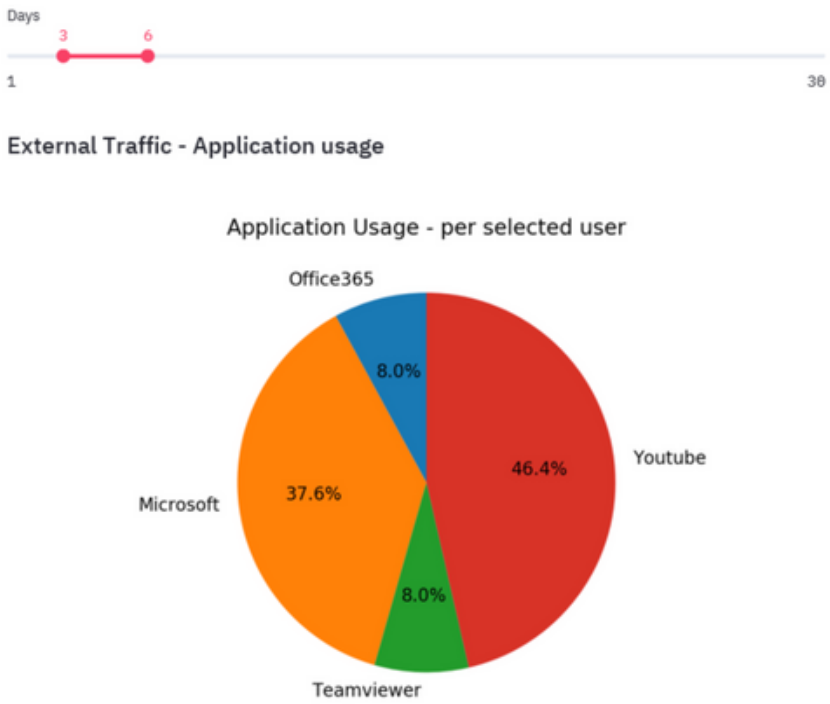
Si las aplicaciones más voluminosas están destinadas a identificar servicios seguros en la nube,

se pueden utilizar con un túnel dividido, lo que reduce la carga en un concentrador VPN. Sin embargo, si las aplicaciones principales son para servicios que son menos seguros o que pueden suponer un riesgo, es más seguro pasarlos a través del túnel VPN. La razón es que otros dispositivos de seguridad de red pueden procesar el tráfico antes de permitir que pase dicho tráfico. A continuación, puede utilizar las políticas de acceso de los firewalls para limitar el acceso a las redes externas.

Usuarios de VPN individuales no conformes con la identidad

En algunos casos, la oleada podría estar asociada con unos pocos usuarios que no cumplen con ciertas políticas. Los módulos y conjuntos de datos que se han utilizado anteriormente se pueden volver a utilizar para identificar los principales usuarios de VPN y las aplicaciones web a las que acceden. Esto puede ayudar en el aislamiento de tales usuarios y observar su efecto en la carga del dispositivo.





En los casos en los que ninguno de los métodos encaja, los administradores deben buscar soluciones de seguridad para terminales como la solución AMP para terminales y la solución Cisco Umbrella para proteger los terminales en redes no protegidas.